



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2016

Le emozioni nel parlato doppiato: uno studio cross-linguistico

Maffia, Marta ; Pellegrino, Elisa

Abstract: The study of vocal emotions poses different methodological problems, especially regarding the style of speech analyzed. If spontaneous speech on the one hand guarantees data authenticity, on the other hand does not allow to control the variables affecting spoken communication. Recited speech allows to control the verbal content of emotional productions and obtain easily comparable data. Dubbed speech represents another feasible alternative. The task of a dubber is to use a different language to vehiculate the same emotional state expressed in the original version, to sound as natural as possible. The present study has twofold objective: firstly it aims to describe the prosodic characteristics of dubbed speech, to evaluate its effectiveness for the cross-linguistic analysis of vocal emotions; secondly it is intended to verify if romance and non-romance languages present similar prosodic patterns in the expression of different emotional states. To reach these goals, sequences taken from the original and the dubbed versions (Italian, French, Spanish, Portuguese and German) of “Titanic” have been the object of perceptive tests and a spectro-acoustic analysis. The strategies used by romance and nonromance languages to obtain synchrony and the acoustic correlates of five perceived basic emotions (anger, joy, sadness, fear and surprise) have been identified and discussed in details.

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: <https://doi.org/10.5167/uzh-127829>
Book Section

Originally published at:

Maffia, Marta; Pellegrino, Elisa (2016). Le emozioni nel parlato doppiato: uno studio cross-linguistico. In: Bardel, Camilla; De Meo, Anna. Parler les langues romanes/Parlare le lingue romanze/Hablar las lenguas romances/Falando línguas românicas. Naples: Il Torcoliere, 111-126.

Parler les langues romanes
Parlare le lingue romanze
Hablar las lenguas romances
Falando línguas românicas

Edited by

Camilla Bardel and Anna De Meo

Naples
2016



Prodotto da

IL TORCOLIERE • *Officine Grafico-Editoriali d'Ateneo*
UNIVERSITÀ DEGLI STUDI DI NAPOLI "L'Orientale"
Settembre 2016

ISBN: 978-88-6719-111-6

MARTA MAFFIA & ELISA PELLEGRINO

LE EMOZIONI NEL PARLATO DOPPIATO: UNO STUDIO CROSS-LINGUISTICO

Abstract

The study of vocal emotions poses different methodological problems, especially regarding the style of speech analyzed. If spontaneous speech on the one hand guarantees data authenticity, on the other hand does not allow to control the variables affecting spoken communication. Recited speech allows to control the verbal content of emotional productions and obtain easily comparable data. Dubbed speech represents another feasible alternative. The task of a dubber is to use a different language to vehiculate the same emotional state expressed in the original version, to sound as natural as possible. The present study has twofold objective: firstly it aims to describe the prosodic characteristics of dubbed speech, to evaluate its effectiveness for the cross-linguistic analysis of vocal emotions; secondly it is intended to verify if romance and non-romance languages present similar prosodic patterns in the expression of different emotional states. To reach these goals, sequences taken from the original and the dubbed versions (Italian, French, Spanish, Portuguese and German) of "Titanic" have been the object of perceptive tests and a spectro-acoustic analysis. The strategies used by romance and nonromance languages to obtain synchrony and the acoustic correlates of five perceived basic emotions (anger, joy, sadness, fear and surprise) have been identified and discussed in details.

1. Introduzione

La manifestazione delle emozioni e la loro percezione sono state e sono tuttora oggetto di numerose indagini in diversi ambiti di ricerca: dalla biologia alla linguistica, dalla psicologia alla fonetica. È stato dimostrato come, nella specie umana, gli stati emotivi siano espressi attraverso una pluralità di canali comunicativi, che includono la voce, l'espressione facciale, il *body language* e il contenuto verbale del messaggio (Ekman, 1982; Poggi & Magno Caldognetto, 2004). Sebbene la discussione sia ancora aperta su quale sia l'apporto di ciascun canale nella trasmissione delle diverse emozioni, è opinione ormai condivisa che i tratti prosodici siano indicatori essenziali per il riconoscimento del parlante e per la decodifica del parlato emotivo (Anolli & Ciceri, 1992; Mozziconacci, 2002; Scherer, Johnstone & Klasmayer, 2003).

L'identificazione dei correlati acustici delle emozioni è stata infatti oggetto di numerosi studi di carattere intra-linguistico. In particolare, sulla base del modello componenziale delle emozioni proposto da Scherer (2003), tali ricerche hanno dimostrato che esiste una correlazione tra la dimensione emotiva dell'attivazione (dimensione che indica il grado di tendenza all'azione da parte del parlante e definita anche *activity* o *arousal*, Cowie et alii, 2001) e la variazione di alcuni parametri acustici (Johnstone & Scherer, 2000; Schröder et alii, 2001; Schröder, 2003). Il grado di attivazione sembra essere positivamente correlato ai valori medi di f_0 , del *range* tonale e dell'intensità nonché della velocità di articolazione. Le emozioni ad alta attivazione, inoltre, come la collera e la gioia (al contrario delle emozioni a bassa attivazione, come la tristezza e il disgusto) sarebbero veicolate anche da una minore percentuale di silenzi sul totale dell'enunciato e una più contenuta durata media delle pause (Trouvain & Barry, 2000). Altro parametro acustico solitamente in correlazione con l'espressione degli stati emotivi, ma che risulta più difficilmente associabile a variazioni in termini di grado di attivazione, è la qualità della voce (Golb & Ni Chasaide, 2003).

Anche studi di carattere cross-linguistico sulla produzione e sulla percezione delle emozioni dalla voce hanno enfatizzato il ruolo delle caratteristiche prosodiche dell'enunciato nell'identificazione di diversi stati emotivi. Sebbene non si neghi l'esistenza delle cosiddette regole di esibizione (*display rules*) identificate per la prima volta da Ekman & Friesen (1975) e specifiche di ciascuna lingua e cultura, studi condotti in prospettiva cross-linguistica hanno ipotizzato che diversi stati emotivi siano veicolati da modelli prosodici universali, condivisi da tutte le lingue (Pfitzinger et al., 2011; Laukka 2004; Kori & Magno Caldognetto, 2003; Yang & Campbell, 2001).

Una delle maggiori sfide poste da queste ricerche sperimentali sul parlato espressivo, sia di tipo intra-linguistico sia cross-linguistico, riguarda la metodologia di raccolta dei dati (Scherer, 2003; Coan & Allen, 2007). È innanzitutto necessario scegliere lo stile di parlato da analizzare: un parlato spontaneo garantisce l'autenticità delle emozioni ma non permette la comparabilità dei dati né il controllo su un largo numero di variabili che possono condizionare la comunicazione parlata (come la presenza di sovrapposizioni, rumori di sottofondo, etc). Nel parlato recitato, al contrario, benché si abbia il vantaggio di poter controllare agevolmente il contenuto testuale delle produzioni, si può incorrere in un duplice rischio: da un lato uno stesso testo potrebbe essere interpretato con coloriture emozionali diverse, dall'altro gli attori potrebbero produrre versioni stereotipate delle emozioni, che difficilmente riflettono quello che naturalmente i parlanti

producono e percepiscono. Altre problematiche di carattere metodologico riguardano la natura dei parlanti coinvolti (attori o *naïve*), il tipo di materiale linguistico sul quale si vuole effettuare l'analisi (parole isolate, *nonsense*, sillabe, enunciati, etc) e, infine, i metodi di induzione degli stati emotivi (video, interviste, etc).

È fuori di dubbio che scelte metodologiche diverse possono condizionare i risultati ottenuti e che, viceversa, i risultati ottenuti attraverso procedure tanto diverse possono essere difficilmente comparabili.

2. Il parlato doppiato

Una tipologia di parlato che consente di superare, almeno in parte, i limiti metodologici evidenziati è il parlato doppiato. Contrariamente a quanto accade nel caso di dati spontanei, nel doppiaggio le variabili della comunicazione sono sotto controllo: il contesto linguistico ed extralinguistico, l'avvicinarsi dei turni e gli interlocutori sono stabiliti nella versione originale del film.

Il doppiatore, inoltre, deve adeguare le proprie battute alla mimica facciale, ai gesti, alla postura e allo sguardo dell'attore che per primo ha girato le scene.

Oltre a consentire la raccolta di dati comparabili dal punto di vista della situazione comunicativa, il parlato doppiato presenta anche il vantaggio di essere meno artificiale rispetto a quello di laboratorio: lo scopo della messa in scena non è quello di fornire "buoni esempi" di diversi stati emotivi, ma di produrre enunciati che risultino il più possibile verosimili.

In questo senso, il doppiatore, nell'interpretare una battuta con una determinata sfumatura emotiva, deve dare la propria voce a un personaggio di fatto già esistente, che ha uno specifico ritmo respiratorio e, soprattutto, una specifica intenzione. Lo scopo del doppiaggio, infatti, è quello di usare una nuova lingua per veicolare lo stesso stato emotivo espresso nella versione originale, al fine che ogni scena sembri naturale e non vi siano incoerenze tra video e audio.

Studi condotti sulla prosodia del parlato doppiato ne hanno messo in evidenza specifiche caratteristiche ritmiche e di velocità di articolazione dovute al fatto che, nell'operazione di sincronizzazione, la lingua di doppiaggio deve necessariamente adattarsi ai tempi della lingua originale. Sono state evidenziate diverse strategie per l'ottenimento della sincronia con il video, che sembrano dipendere dalle caratteristiche ritmiche della lingua

originale e delle lingue di doppiaggio esaminate. In particolare, nel caso in cui la lingua originale era l'inglese, la strategia riscontrata è stata un'accelerazione nell'articolazione del parlato doppiato in italiano; nel caso inverso, in cui la lingua originale era l'italiano, piuttosto che una decelerazione della velocità di articolazione, si è riscontrato nel doppiaggio inglese un aumento delle pause silenti (Pettorino & Giannini, 2000; Pettorino, 2000; Pettorino & Vitagliano, 2003).

Pochi studi, infine, si sono concentrati sulle caratteristiche prosodiche delle emozioni nel parlato doppiato, sia in prospettiva diacronica (Vitagliano, Giannini & Pettorino, 2008) sia in prospettiva cross-linguistica (Pettorino, 2004; Braun & Katerbow, 2005; Braun & Oba, 2007). I risultati di tali studi hanno dimostrato che, seppure nei limiti temporali imposti dalla versione originale, esistono anche nel parlato doppiato variazioni ritmiche e intonative, a volte condivise, a volte specifiche di una determinata lingua, legate all'espressione di diversi stati emotivi.

3. Obiettivi

Il presente studio si propone in primo luogo di esaminare le diverse strategie utilizzate da lingue appartenenti a diversi gruppi ritmici per ottenere la sincronia nel doppiaggio (partendo dall'inglese come lingua originale).

Ulteriore scopo della ricerca è quello di confrontare, in chiave cross-linguistica, i correlati acustici di specifici stati emotivi all'interno di un corpus di parlato doppiato composto da lingue romanze e non romanze.

4. Materiali e metodo

Per raggiungere tali obiettivi, è stata effettuata un'analisi su porzioni di parlato tratte dal film *Titanic* del 1997 nella versione originale in inglese e nelle seguenti lingue di doppiaggio: francese, italiano, portoghese, spagnolo e tedesco. In primo luogo sono state individuate alcune sequenze del film che fossero emotivamente più connotate.

Successivamente si è proceduto con la selezione di singole scene, avendo l'accortezza di identificare porzioni in cui vi fosse un unico parlante (evitando quindi le sovrapposizioni) e in cui non fosse presente musica di sottofondo (che avrebbe inficiato o almeno complicato la successiva analisi acustica). In totale sono state selezionate ed estratte 156 scene (26 per ciascuna lingua).

4.1. Test percettivi

Le sequenze estratte dal film *Titanic* sono state successivamente oggetto di test percettivi, volti a selezionare le scene che veicolassero lo stesso stato emotivo in tutte le lingue considerate. I test, in lingua originale e nelle diverse lingue di doppiaggio, sono stati somministrati online attraverso la piattaforma *Questbase*.

Hanno partecipato ai test 25 parlanti nativi di ciascuna lingua esaminata, per un totale di 150 partecipanti, tutti studenti o docenti all'Università degli Studi di Napoli "L'Orientale" e all'Università di Cambridge, con un'età media di 32 anni.

Agli ascoltatori è stato chiesto di osservare le 26 scene nella propria lingua materna (proposte in ordine casuale) e di individuare lo stato emotivo espresso dal parlante in ciascuna sequenza. La scelta poteva essere effettuata tra le sei opzioni proposte, corrispondenti alle sei emozioni di base (Ekman, 1992), oppure aggiungendo una nuova etichetta che fosse ritenuta più appropriata.

Per l'individuazione delle etichette emotive nelle diverse lingue si è fatto riferimento, per il francese, l'inglese, l'italiano e il tedesco, alle traduzioni proposte da Scherer (1988) mentre per lo spagnolo e il portoghese sono stati coinvolti parlanti nativi delle due lingue esperti di linguistica. Inoltre, per ciascuna scena, è stato chiesto ai partecipanti al test di valutare anche il grado di intensità emotiva su una scala da 1 (bassa intensità) a 3 (alta intensità).

4.2. Il corpus

I risultati dei test percettivi hanno permesso di selezionare le scene in cui la stessa emozione fosse riconosciuta in tutte le lingue da almeno il 60% degli ascoltatori e in cui lo scarto tra l'etichetta con la più alta percentuale di riconoscimento e le altre etichette fosse significativo.

Il corpus così ottenuto è composto da 14 sequenze di parlato per ciascuna lingua (84 in totale), rappresentative di cinque diversi stati emotivi: collera, gioia, paura, sorpresa e tristezza. Per ciascuna emozione il corpus comprende sequenze prodotte da almeno una voce femminile e una maschile, ad eccezione della tristezza, per cui sono presenti esclusivamente voci femminili.

4.3. L'analisi spettroacustica

L'analisi spettroacustica è stata condotta manualmente, attraverso l'ausilio dei software *Wavesurfer* e *Praat*. Per ciascuna porzione di parlato in ciascuna lingua sono state effettuate le seguenti misure:

- numero e durata delle catene foniche;
- numero e durata delle sillabe;
- numero e durata delle pause silenti;
- durata degli *affect bursts*;
- valore minimo e massimo di f_0 per ciascuna catena fonica.

Nel presente studio si utilizza il termine *burst* per fare riferimento, secondo la definizione di Scherer (1994), alle "very brief, discrete, nonverbal expressions of affect", la cui quantità e la cui tipologia sono stati definiti componenti importanti nella manifestazione vocale delle diverse emozioni.

Le misure effettuate hanno permesso di ricavare quegli indici prosodici che in letteratura sono generalmente associati all'espressione delle emozioni:

- velocità di Articolazione - VdA (sill/sec);
- fluenza (sill/n. di catene foniche);
- durata media delle pause silenti;
- composizione dell'enunciato (% di silenzi, sillabe e *bursts*);
- *range* tonale (calcolato in Herzt e successivamente normalizzato in semitoni).

Al fine di rendere comparabili i dati dei diversi parlanti e nelle diverse lingue, per ogni enunciato il valore minimo e massimo di f_0 sono stati messi in relazione al più basso valore di f_0 (f_0 floor) raggiunto da ciascun parlante nell'intero corpus. Attraverso questa procedura è stato possibile ricavare il registro tonale di ciascuna emozione in inglese e nelle cinque lingue di doppiaggio.

Per quanto riguarda la qualità della voce, rimandiamo a una fase successiva del lavoro l'analisi acustica dei parametri che la caratterizzano. Nel presente studio si proporranno esclusivamente considerazioni basate su un'analisi percettiva della qualità della voce in ciascuna scena selezionata per ciascuna lingua, condotta da un ascoltatore esperto.

5. Risultati

5.1. Strategie di doppiaggio

Come si può osservare in tabella 1, in tutte le versioni doppiate, la durata totale delle porzioni selezionate non presenta differenze significative con la

versione originale, ad eccezione della versione portoghese, che risulta di circa 5 secondi più lunga.

Quello che differisce notevolmente è invece il numero di sillabe, che è più alto in tutte le lingue di doppiaggio, con un aumento che raggiunge circa il 50% nel caso dell'italiano e che risulta invece meno accentuato nella versione francese.

Al contrario, il numero delle catene foniche subisce un decremento in tutte le versioni doppiate, testimoniato dal minor numero di pause silenti. Tali pause, inoltre, diminuiscono nelle lingue di doppiaggio anche nella durata media, con l'eccezione della lingua portoghese.

Sulla base dei dati finora presentati è interessante notare come il francese (piuttosto che il tedesco, che appartiene allo stesso gruppo ritmico della lingua originale) risulti la lingua che si comporta in maniera più simile all'inglese, in termini di durate globali.

	Inglese	Italiano	Francese	Spagnolo	Portoghese	Tedesco
Durata tot. (s)	169,6	170,3	169,1	167,3	176,9	171,5
N. sillabe	555	763	656	730	664	677
N. pause silenti	58	53	56	51	53	54
Durata media delle pause (s)	0,84	0,78	0,79	0,74	0,86	0,73

Tab. 1: Durata totale, numero di sillabe, numero e durata media dei silenzi nelle diverse lingue.

Questa osservazione si conferma se esaminiamo i dati sulla composizione dell'enunciato (Fig. 1).

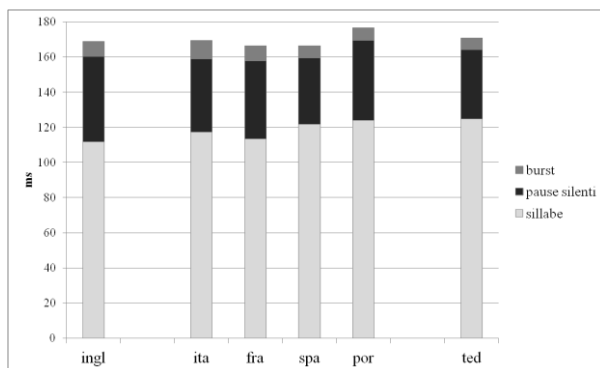


Fig. 1: Composizione del parlato nelle diverse lingue.

Mentre il francese presenta valori molto vicini alla versione inglese originale, nelle versioni italiana, spagnola e tedesca la percentuale di silenzio è ridotta, a vantaggio del tempo occupato dalle sillabe.

Nella versione portoghese, invece, sebbene vi sia un aumento nella percentuale di sillabe prodotte, non si riscontra una riduzione drastica dei silenzi.

Nelle versioni spagnola e tedesca, inoltre, anche la percentuale di *bursts* decresce, cosa che non accade per l'italiano.

In tutte le lingue di doppiaggio, come si può vedere in figura 2, si assiste a un aumento della velocità di articolazione rispetto alla versione originale, meno accentuato nel caso del portoghese, in cui, come è stato già detto, la più alta percentuale di sillabe non corrisponde a una diminuzione della durata dei silenzi.

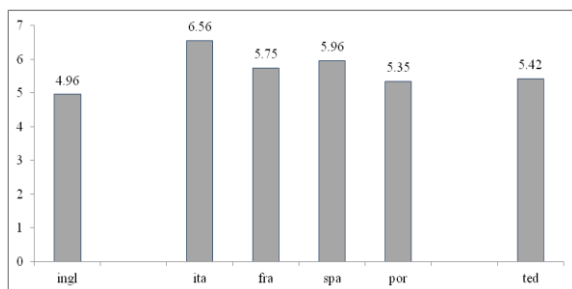


Fig. 2: Velocità di articolazione nelle diverse lingue.

Le lingue romanze esaminate, quindi, per ottenere la sincronia con i tempi della lingua originale, utilizzano (ad eccezione del portoghese) una

combinazione delle strategie riscontrate in letteratura (aumento della VdA e diminuzione delle pause silenti). Non sembra esserci, inoltre, in termini di strategie di doppiaggio, una distinzione peculiare tra lingue romanze, tutte a tendenza isosillabica, e l'unica lingua non romanza del corpus a tendenza isoaccentuale, il tedesco.

5.2. Correlati acustici delle emozioni

È stato già constatato che globalmente la velocità di articolazione aumenta in tutte le versioni doppiate, per effetto della compressione di un maggior numero di sillabe in tempi più o meno simili. Se si osservano i dati su questo indice relativi ai diversi stati emotivi (Fig. 3) si può notare che le variazioni risultano più o meno accentuate, a seconda delle singole emozioni riconosciute.

La prima considerazione da fare è che nella versione originale sembra esserci meno variazione interna in termini di VdA rispetto alle versioni doppiate, in cui le differenze tra i diversi stati emotivi risultano più accentuate.

Le emozioni veicolate dai più alti valori di VdA, coerentemente con i dati già presenti in letteratura, risultano la collera (per tutte le lingue considerate) e la gioia (ad eccezione della versione originale inglese), entrambe emozioni ad alta attivazione. I valori più bassi di VdA sono, invece, quelli relativi alla tristezza (a bassa attivazione) e alla sorpresa. Quest'ultimo dato sulla sorpresa, più che al grado di attivazione emotiva (anche la sorpresa è un'emozione ad alta attivazione) sembra essere condizionato dall'avanzata età della parlante donna (l'anziana Rose del film), in accordo con i risultati degli studi condotti sui cambiamenti della voce legati all'età (Pettorino, Pellegrino & Maffia, 2014).

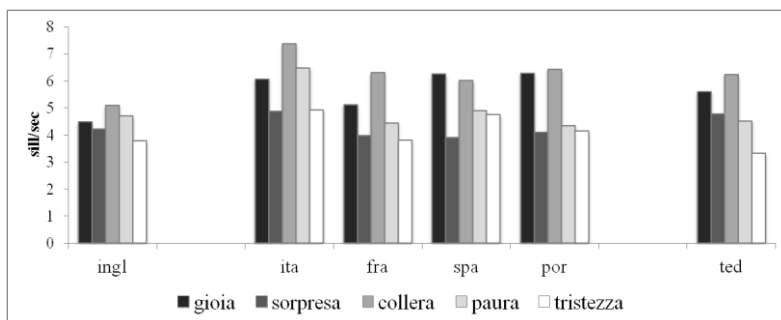


Fig. 3: Velocità di articolazione nelle diverse lingue e nei diversi stati emotivi riconosciuti.

Riguardo la fluenza, è stato già osservato come uno degli effetti del doppiaggio sia la diminuzione del numero di catene foniche, legata al minor numero di silenzi nelle versioni doppiate. Ciò significa, quindi, che le catene foniche nelle lingue di doppiaggio dovrebbero risultare più lunghe di quelle della versione originale. I dati sulla fluenza ripartiti per emozione (Fig. 4) ci dicono, però, che l'aumento nella durata delle catene foniche è condizionato anche dallo stato emotivo espresso dal parlante. Anche in questo caso, un aumento maggiore si riscontra, infatti, nel caso della gioia e della collera, in quasi tutte le lingue di doppiaggio.

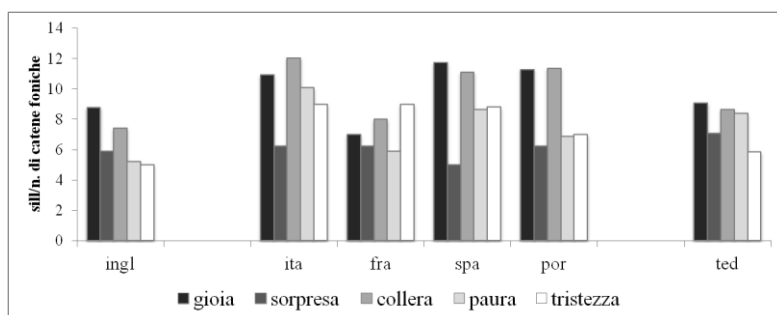


Fig. 4: Valori della fluenza nelle diverse lingue e nei diversi stati emotivi riconosciuti.

Anche la durata delle pause silenti, che subisce un decremento in tutte le versioni doppiate, sembra però essere condizionata dai diversi stati emotivi espressi. Se si osserva il grafico in Fig. 5 si può notare infatti che, nelle porzioni riconosciute come tristi, in accordo con i dati presenti in letteratura, le pause risultano in media più lunghe.

Anche in questo caso, i valori insoliti relativi alla sorpresa possono essere condizionati dall'età della parlante donna.

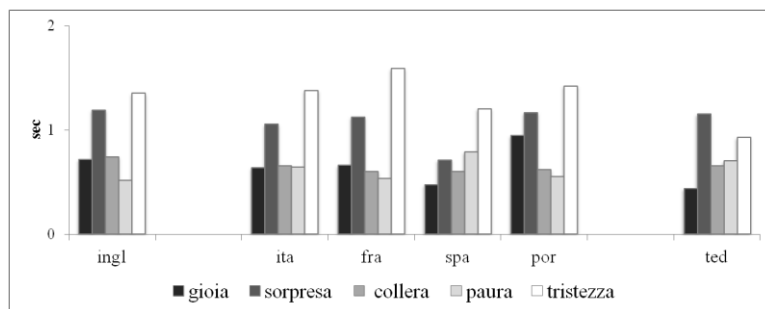


Fig. 5: Durata media delle pause silenti nelle diverse lingue e nei diversi stati emotivi riconosciuti.

Per quanto riguarda i dati sulla composizione dell'enunciato, non vi sono differenze significative tra le diverse lingue. Si può notare, a conferma e completamento di quanto detto sulla durata media delle pause silenziose, che si riscontra, come tendenza comune, una maggiore percentuale di silenzi nel caso della tristezza e della sorpresa. Come atteso, i *bursts* sono presenti in tutte le lingue, quasi esclusivamente nelle porzioni riconosciute come gioiose. È comprensibilmente molto raro che essi possano essere introdotti nelle versioni doppiate nel caso in cui siano assenti in lingua originale o, viceversa, che *bursts* originali siano doppiati con parlato di altro tipo, a causa soprattutto delle limitazioni visive (espressioni facciali specifiche legate ai *bursts*) e ai ritmi di respirazione dell'attore nella scena girata.

Nel caso del *range* tonale e del registro, i dati risultano più variati rispetto a quelli precedentemente presentati (Fig. 6). Al di là delle sequenze riconosciute come gioiose e tristi, caratterizzate (in parziale disaccordo con i dati presenti in letteratura) da un registro basso e un'escursione tonale contenuta, per gli altri stati emotivi è difficile individuare delle tendenze comuni, se non un ampio *range* tonale nel caso della sorpresa e un registro in media più alto nel caso della paura.

Infine, le osservazioni percettive condotte da un ascoltatore esperto, sulla base dei *settings* laringali identificati da Laver (1980), riguardo la qualità della voce in ciascuna porzione di parlato e in ciascuna lingua analizzata, hanno evidenziato e confermato che non esiste un'associazione univoca tra tale parametro e i diversi stati emotivi. In particolare, la stessa qualità della voce può veicolare diversi stati emotivi, anche con diversi gradi di attivazione: ad esempio una voce sospirata è stata identificata sia nelle porzioni riconosciute come tristi sia in quelle colleriche.

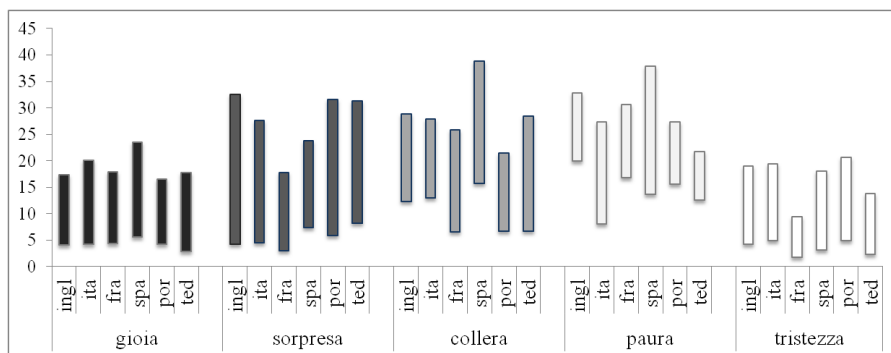


Fig. 6: Range tonale e registro nelle diverse lingue e nei diversi stati emotivi riconosciuti.

5.3. Note sull'intensità emotiva

Riguardo la valutazione percettiva dell'intensità emotiva, è stato constatato che nel nostro corpus esiste una correlazione positiva ($r: 0,41$) tra la percentuale di riconoscimento condiviso e la percentuale di giudizi di alta intensità emotiva: più un'emozione è stata definita intensa, maggiore è stato il consenso nel suo riconoscimento. Tali dati confermano quelli già presenti in letteratura sul ruolo dell'intensità nel processo di decodifica delle emozioni (Juslin & Laukka, 2001).

Se si esaminano i dati relativi all'intensità percepita nei diversi stati emotivi, è interessante notare che, indipendentemente dalla lingua di doppiaggio e dal retroterra culturale degli ascoltatori, le emozioni che hanno ottenuto le più alte percentuali di giudizi di alta intensità sono la paura, la collera e, in alcuni casi, la tristezza (Fig. 7).

Tali stati emotivi risultano anche tra quelli più caratterizzati dal punto di vista prosodico.

La sorpresa, invece, è l'emozione che ha ricevuto in tutte le lingue le più basse percentuali di giudizi di alta intensità, probabilmente a causa delle insolite caratteristiche prosodiche già evidenziate, legate all'età della parlante donna.

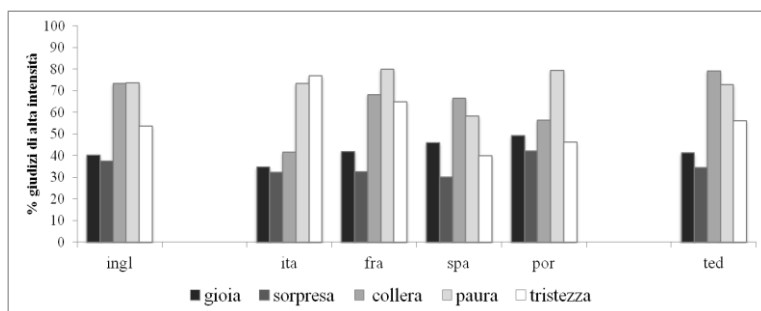


Fig. 7: Giudizi di alta intensità nelle diverse lingue e nei diversi stati emotivi riconosciuti.

6. Conclusioni

L'analisi condotta sul parlato doppiato in diverse lingue romanze e non romanze, con l'inglese come lingua originale, ha messo in evidenza che le due strategie riscontrate in letteratura per ottenere la sincronia (aumento della

VdA e diminuzione della durata delle pause silenti) non sono necessariamente in alternativa l'una all'altra, ma possono anche essere combinate. Inoltre, non è stato riscontrato un comportamento diverso nella lingua tedesca (unica non romanza e a tendenza isoaccentuale) rispetto alle altre lingue di doppiaggio considerate (tutte romanze e a tendenza isosillabica).

Per quanto riguarda invece l'analisi dell'espressione vocale delle emozioni, è da sottolineare innanzitutto che i risultati dei test percettivi sembrano confermare l'efficacia dell'uso del parlato doppiato in ricerche di carattere cross-linguistico: le stesse scene in lingue diverse sono state effettivamente riconosciute, con un notevole grado di omogeneità, come rappresentative degli stessi stati emotivi.

L'analisi spettroacustica, inoltre, ha permesso di evidenziare che, seppure il parlato doppiato presenta delle proprie, intrinseche caratteristiche prosodiche, è possibile riscontrare variazioni acustiche legate all'espressione di diversi stati emotivi.

A conferma dei risultati di studi precedenti condotti in prospettiva cross-linguistica, e che hanno ipotizzato l'universalità di certe variazioni prosodiche nel parlato emotivo, all'interno del corpus, sono state individuate alcune tendenze comuni a tutte le lingue esaminate relative alle specifiche emozioni di base riconosciute:

- collera e gioia sono veicolate da alti valori di VdA e di fluenza;
- collera e paura presentano le maggiori percentuali di sillabe e un registro tonale in media più alto;
- la sorpresa è caratterizzata da un ampio *range* tonale e lunghe pause silenti (età del parlante);
- la tristezza è veicolata da bassi valori di VdA, una maggiore percentuale di silenzio, un registro basso e un *range* tonale ristretto.

In generale è possibile constatare che, in questo particolare tipo di corpus di parlato doppiato, se le caratteristiche ritmiche possono essere soggette a variazioni minime nelle diverse lingue, condizionate dai tempi e dalle pause della versione originale, è invece negli aspetti intonativi che sembra riscontrarsi una maggiore variabilità.

7. Prospettive future

Al fine di rendere i risultati di tale studio più generalizzabili e di approfondire l'analisi delle caratteristiche prosodiche delle emozioni nel parlato doppiato, si rende sicuramente necessario un ampliamento del

corpus, sia in termini di numero di scene e di quantità di parlato analizzato sia di termini di lingue considerate.

Come già detto in precedenza, inoltre, gli aspetti prosodici rappresentano solo uno dei canali attraverso i quali gli stati emotivi sono veicolati. Un'analisi più accurata delle emozioni nel parlato doppiato dovrebbe perciò tenere conto di altre variabili quali: il processo di traduzione e gli aspetti testuali, il contenuto non verbale delle scene (ambientazione, linguaggio del corpo e espressioni facciali degli attori) e il tipo di ripresa considerata (primo piano o figura intera).

BIBLIOGRAFIA

- ANOLLI L. & CICERI R. (1992), *La voce delle emozioni*, Milano, Franco Angeli.
- BRAUN A. & KATERBOW M. (2005), "Emotions in Dubbed Speech: An intercultural approach with respect to F0", in *Proceedings of the Conference of the International Speech Communication Association*, 4-8 settembre 2005, Lisbona, Portogallo, 521-524.
- BRAUN A. & OBA R. (2007), "Speaking Tempo in Emotional Speech – a Cross-Cultural Study Using Dubbed Speech", in *Proceedings 16th International Conference on Phonetic Sciences*, 6-10 agosto 2007, Saarbrücken, Germania, 77-82.
- COAN, J. A. & ALLEN, J. J. B. (a cura di) (2007), *The Handbook of Emotion Elicitation and Assessment*, Oxford, Oxford University Press.
- COWIE R., DOUGLAS-COWIE E., TSAPATSOU LIS N., VOTSIS G., KOLLIAS S., FELLE NZ W. & TAYLOR J. (2001), "Emotion recognition in human-computer interaction", in *IEEE Signal Processing Magazine*, Vol. 18, 32-80.
- EKMAN P. & FRIESEN W. V. (1975), *Unmasking the face*, Englewood Cliffs, N.J., Prentice-Hall.
- EKMAN P. (1982), *Emotion in human face*, Cambridge Mass., Cambridge University Press.
- EKMAN P. (1992), "An argument for basic emotions", in *Cognition and Emotion*, Vol. 6, 169-200.
- GOBL C. & NÍ CHASAIDE A. (2003), "The role of voice quality in communicating emotion, mood and attitude", in *Speech Communication*, Vol. 40, no. 1-2, 189-212.
- JOHNSTONE T. & SCHERER K. R. (2000), "Vocal communication in Emotion", in M. Lewis, J. Haviland (a cura di) *Handbook of Emotion*, New York, The Guilford Press, 220-235.

- JUSLIN P. N. & LAUKKA, P. (2001), "Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion", in *Emotion*, Vol. 1, 381-412.
- KORI S. & MAGNO CALDOGNETTO E. (2003), "La caratterizzazione fonetica delle emozioni: primi dati di uno studio cross-linguistico italiano-giapponese", in P. Cosi, E. Magno Caldognetto, A. Zamboni (a cura di) *Voce-Canto-Parlato. Studi in onore di Franco Ferrero*, Padova, Unipress, 187-200.
- LAUKKA P. (2004), *Vocal expression of emotion: discrete-emotions and dimensional accounts*, Tesi di Dottorato, Uppsala University.
- LAVER J. (1980), *The phonetic description of voice quality*, Cambridge, Cambridge University Press.
- MOZZICONACCI S. (2002), "Prosody and Emotions", in *Proceedings of Speech Prosody 2002*, Aix-en-Provence, 11-13 aprile 2002, 1-9.
- PETTORINO M. (2000), "Alcune osservazioni sul doppiaggio cinematografico", in *Atti del XXVIII Convegno Nazionale AIA*, 10-13 giugno 2000, Trani.
- PETTORINO M. (2004), "Inglese italiano e giapponese: analisi dei correlati acustici delle emozioni nel parlato cinematografico", in *Atti del 1° Convegno Nazionale GSCP 2004*, Padova, 30 ottobre - 1 novembre 2004, Napoli, Liguori, 45-57.
- PETTORINO M. & GIANNINI A. (2000), "My name is Bond, James Bond / Il mio nome è Bond, James Bond", in D. Locchi (a cura di) *Il parlante e la sua lingua. Atti delle X Giornate di Studio del Gruppo di Fonetica Sperimentale (GFS)*, 13-15 dicembre 1999, Napoli, Il Torcoliere, 169-176.
- PETTORINO M., PELLEGRINO E. & MAFFIA M. (2014), "'Young' and 'Old' Voice: the prosodic auto-transplantation technique for speaker's age recognition", in N. Campbell, D. Gibbon, D. Hirst (a cura di) *Social and Linguistic Speech Prosody*, 20-23 maggio 2014, Dublino, Irlanda, 135-139.
- PETTORINO M. & VITAGLIANO I. (2003), "Prosodic Characteristics of Dubbed Speech", in *Proceedings of the XV International Congress of Phonetic Sciences*, 3-9 agosto 2003, Barcelona, 2865-2868.
- PFITZINGER H. R., AMIR N., MIXDORFF H. & BÖSEL J. (2011), "Cross-language perception of Hebrew and German authentic emotional speech", in *Proceedings of ICPHS2011*, 17-21 agosto 2011, Hong Kong, 1586-1589.
- POGGI I. & MAGNO CALDOGNETTO E. (2004), "Il parlato emotivo. Aspetti cognitivi, linguistici e fonetici", in F. Albano Leoni, F. Cutugno, M. Pettorino, R. Savy (a cura di), *Il parlato Italiano*, Napoli, M. D'Auria Editore, CD-Rom.
- SCHERER K. R. (1988), *Facets of emotion: Recent research*, Hillsdale, NJ, Erlbaum.

- SCHERER K. R. (1994), "Affect Bursts", in S. H. M. van Goozen, N. E. van de Poll, J. A. Sergeant (a cura di), *Emotions*, Hillsdale, NJ, Lawrence Erlbaum, 161-193.
- SCHERER K. R. (2003), "Vocal Communication of Emotion: a Review of Research Paradigm", in *Speech Communication*, Vol. 40, 227-256.
- SCHERER K. R., JOHNSTON T. & KLASMEYER G. (2003), "Vocal expression of emotion", in R. J. Davidson, K. R. Scherer & H. H. Goldsmith (a cura di), *Handbook of Affective Science*, New York, Oxford University Press, 433-456.
- SCHRÖDER M. (2003), *Speech and Emotion research. An overview of Research Frameworks and a Dimensional Approach to Emotional Speech Synthesis*, Tesi di dottorato, Universität der Saarlandes.
- SCHRÖDER M., COWIE R., DOUGLAS-COWIE E., WESTERDIJK M. & GIELEN S. (2001), "Acoustic Correlates of Emotion Dimensions in View of Speech Synthesis", in *Proceedings of the EUROSPEECH 2001*, Aalborg, Danimarca, 3-7 settembre 2001, Vol. 1, 87-90.
- TROUVAIN J. & BARRY W. J. (2000), "The prosody of excitement in horse race commentaries", in *Proceedings of the ISCA Workshop on Speech and Emotion*, 5-7 settembre 2000, Newcastle, Irlanda del Nord, 86-91.
- VITAGLIANO I., GIANNINI A. & PETTORINO M. (2008), "Le emozioni nel tempo", in *Fenomeni di intensità nell'italiano parlato*, Giornate di Studio PRIN 2006, Università del Salento, Lecce.
- YANG L. & CAMPBELL N. (2001), "Linking form to meaning: the expression and recognition of emotions through prosody", in *Proceedings of the 4th ISCA Workshop on Speech Synthesis*, 29 agosto - 1 settembre 2001. Perthshire, Scozia.